PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC, **121**:1429–1439, 2009 December

The TAOS Project: High-Speed Crowded Field Aperture Photometry

Z.-W. ZHANG,[1] D.-W. KIM,[2] J.-H. WANG,[3] M. J. LEHNER,[3,4,5] W. P. CHEN,[1] Y.-I. BYUN,[2] C. ALCOCK,[5] T. AXELROD,[6] F. B. BIANCO,[4,5] N. K. COEHLO,[7] K. H. COOK,[8] R. DAVE,[9] I. DE PATER,[10] J. GIAMMARCO,[11] S.-K. KING,[3] T. LEE,[3] H.-C. LIN,[1] S. L. MARSHALL,[12,8] R. PORRATA,[13] P. PROTOPAPAS,[9,5] J. A. RICE,[7] M. E. SCHWAMB,[14] S.-Y. WANG,[3] AND C.-Y. WEN[3]

Received 2009 September 20; accepted 2009 October 22; published 2009 November 16

ABSTRACT. We have devised an aperture photometry pipeline for data reduction of image data from the Taiwanese-American Occultation Survey (TAOS). The photometry pipeline has high computational performance, and is capable of real-time photometric reduction of images containing up to 1000 stars, within the sampling rate of 5 Hz. The pipeline is optimized for both speed and signal-to-noise performance, and in the latter category it performs nearly as well as *DAOPHOT*. This paper provides a detailed description of the TAOS aperture photometry pipeline.

1. INTRODUCTION

The Taiwanese American Occultation Survey (TAOS) is conducting a census of the small objects (~1 km diameter) beyond Neptune. The observational strategy is to search for chance stellar occultations by these small objects (Bailey 1976; Roques et al. 1987, 2009). In order to observe these occultation events, TAOS operates an array of four robotic telescopes with fast optics (f/1.9), each equipped with a 2048×2052 CCD camera, which provides a field-of-view (FOV) of about 3 square degrees. The TAOS system is described in detail in Lehner et al. (2009). With this wide FOV, TAOS can monitor as many as 1000 stars simultaneously with all four telescopes. The typical duration of an occultation event by a kilometer-sized object in the Kuiper Belt is about 0.2 s at opposition. To achieve sensitivity to such short-duration events, TAOS operates its CCD cameras in an innovative way to achieve a sampling rate of 5 Hz. Analysis of such images presents a number of challenges (described in the next section), and we have therefore developed a customized photometry package to process them.

The existing CCD photometry packages such as IRAF *apphot*, *DAOPHOT* (Stetson 1987), and *SExtractor* (Bertin & Arnouts 1996) are able to process TAOS image data. However, these packages either provide low signal-to-noise light curves or run too slowly for real-time analysis. Comparisons of the TAOS photometry with these packages will be discussed in § 6.

1.1. TAOS High-Speed Imaging: Zipper Mode

To achieve sampling rates of 5 Hz by using a conventional CCD camera, we have developed a novel CCD readout scheme called *zipper mode*. Zipper-mode operation is described in detail in Lehner et al. (2009). To summarize, when reading out in zipper-mode, the telescope tracks the sky and stars are imaged onto the focal plane. After 105 ms, a *row block* of 76 rows is read out, while the shutter is kept open. The readout operation takes approximately 95 ms. This is followed by another exposure of 105 ms, then another row block is read out. This process is repeated for a long period of time, typically 90 minutes. Note that the exposure and readout time add up to 200 ms, giving us a readout cadence of 5 Hz. Every 32 row blocks that are read out are written to a single FITS[15] file.

Note that each 105 ms exposure measures flux from the entire field. When the first row block is read out, the remaining

[1] Institute of Astronomy, National Central University, No. 300, Jhongda Rd, Jhongli City, Taoyuan County 320, Taiwan; zwzhang@asiaa.sinica.edu.tw.
[2] Department of Astronomy, Yonsei University, 134 Shinchon, Seoul 120-749, Korea.
[3] Institute of Astronomy and Astrophysics, Academia Sinica. P.O. Box 23-141, Taipei 106, Taiwan.
[4] Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104.
[5] Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138.
[6] Steward Observatory, Tucson, AZ 85721.
[7] Department of Statistics, University of California at Berkeley, Berkeley, CA 94720.
[8] Institute for Geophysics and Planetary Physics, Lawrence Livermore National Laboratory, Livermore, CA 94550.
[9] Initiative in Innovative Computing, Harvard University, Cambridge, MA 02138.
[10] Department of Astronomy, University of California at Berkeley, Berkeley CA 94720.
[11] Department of Astronomy and Physics, Eastern University, Saint Davids, PA 19087.
[12] Kavli Institute for Particle Astrophysics and Cosmology, Menlo Park, CA 94025.
[13] Department of Physics, University of California at Berkeley, Berkeley, CA 94270.
[14] Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125.

[15] Flexible Image Transport System. See http://heasarc.gsfc.nasa.gov/docs/heasarc/fits.html.

electrons on the CCD chip are shifted by 76 rows. When the next row block is read out after the second exposure, electrons from the most recent exposure as well as electrons from the second 76 rows of the first exposure are read out. After $2052/76 = 27$ row blocks have been read out, all of the electrons from the entire first exposure have been read out into the first 27 row blocks. After the next readout, all of the electrons from the second exposure have been read out. Once the entire first exposure has been read out, every subsequent readout step will read out electrons from across the entire CCD added together in a single row block.

Zipper-mode readout is illustrated schematically in Figure 1. Placing the row blocks end to end is the equivalent of stacking the images from each exposure with an offset of 76 rows between each one. In the figure, a series of high-cadence images is offset by one row block after each exposure, and stacked to form a series of row blocks placed end to end. The series of images depicted is collected after the entire first exposure has been read out, and therefore electrons collected across the entire CCD chip are contained in each row block. Each star in the field is thus present in each row block.

Zipper-mode images have several features which we needed to take into consideration when designing the photometric reduction package:

1. Row blocks comprise $2148 \times 76 = 163,248$ pixels (including a 50 pixel of overscan region for each of the two amplifiers on the CCD) of image data, and are read out at a rate of 5 Hz. At 2 bytes pixel^{-1}, this corresponds to a data rate of over 1.6 MB s^{-1} for each telescope. A 90 minute zipper-mode data run will produce nearly 35 GB of data from all four telescopes.

2. Each row block contains all stars in the field. Thus the row block images can be very crowded.

3. Photoelectrons from stars in each exposure are located in different row blocks (see the black dots on the gray stare-mode image in Fig. 1). Two (or more) stars from different places in the field may thus become close neighbors in the row block (see Fig. 2). Moreover, such stars come from different epochs in the exposure series. This gives rise to a possible relative motion between the stars when the telescope has a temporal variation in the tracking error (see Lehner et al. 2009, for a detailed discussion). A variable tracking error as large as 0.2 pixels due to the telescope motor controllers is evident in much of the image data, and sudden movements (~0.5 pixels) due to strong wind gusts can produce relative motions between stars. A similar issue is image rotation caused by the imperfect polar alignment, because of which a star at a corner of the field could move by 1–2 pixels over 90 minutes of observation.

4. During the readout stage of zipper-mode operation, it takes just over 1 ms to read out a single row. Photons from stars continue to be collected during this time, and will create streaks in the image as the electrons are transferred (see Fig. 2).

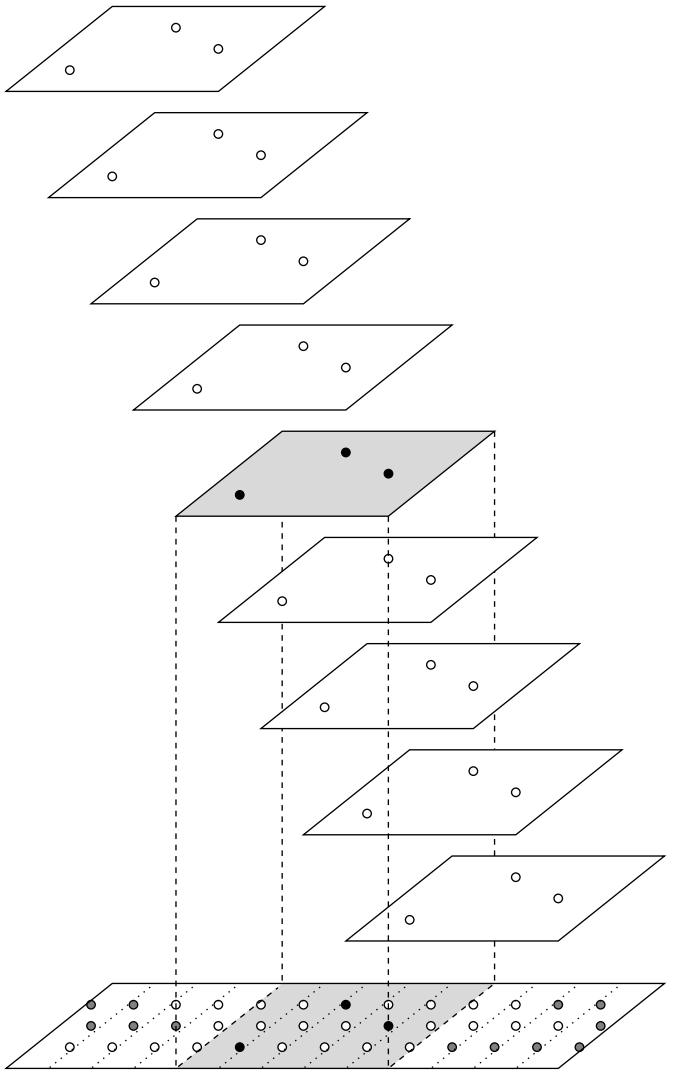

FIG. 1.—Illustration of zipper-mode readout. The top of the diagram shows a sequence of stare-mode images, offset by one row block (*circles* indicate three stars in the field). The bottom shows the resulting ensemble of row blocks (separated by the *dashed lines*). *Open circles* indicate stars shown in the top diagrams, *gray circles* indicate stars from epochs not shown at the top of the diagram. The gray stare-mode image (stars depicted as *filled black circles*) shows up in the row block series as indicated. Each star in the field appears in each row block, but the photons from each star were collected at different epochs.

We have elected to use aperture photometry in order to achieve a sufficiently high data analysis rate to reduce the TAOS image data in real time using a small cluster of computers. To handle the crowdedness of the fields, we have used Fractional Pixel Aperture Photometry (FPAP) with square apertures (described in § 2), with an optimized aperture size for each star (discussed in § 3.3). We have developed a simple algorithm to simultaneously subtract the image streaks and sky background, which is described in § 4.1. Finally, in order to accurately handle time-dependent centroid shifts in the images, we

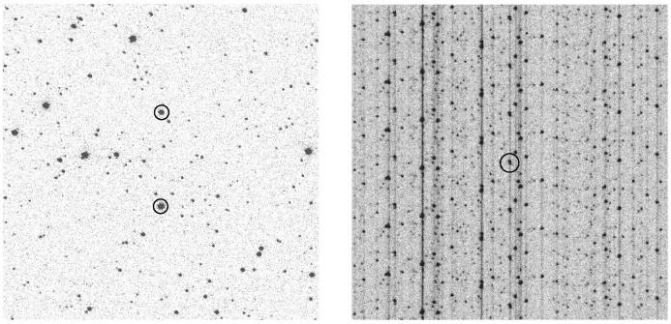

FIG. 2.—*Left panel:* Subimage from a stare-mode image. *Right panel:* Series of zipper-mode row block images, placed end to end, of the same field. Row blocks are placed together in a time sequence with time in the $-y$ direction. The two stars circled in the left panel become close neighbors in the zipper-mode images. The pair of stars is circled in the right panel.

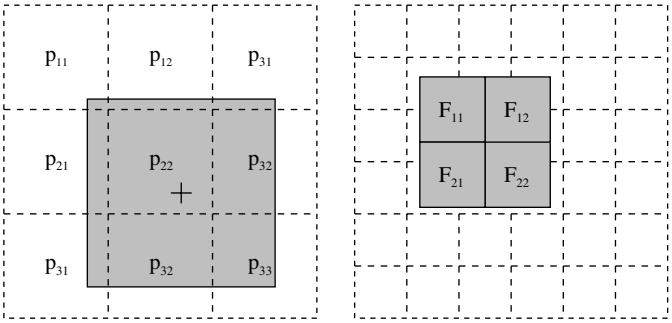

FIG. 3.—*Left panel:* Illustration of fractional pixel aperture photometry with a square aperture. Pixel boundaries are indicated by the dashed lines and are labeled p_{ij}, where i is the row and j is the column. The aperture is indicated with the gray square and is centered on the cross. *Right panel:* Illustration of centroid-finding method. *Dotted lines* indicate pixel boundaries, and *gray squares* indicate fractional pixel apertures. See text for details of algorithm.

have designed our aperture mask on the basis of *exposure* rather than row block. This is discussed in § 4.

2. FRACTIONAL PIXEL APERTURE PHOTOMETRY AND CENTROID-FINDING ALGORITHM

FPAP and our centroid-finding algorithm are used both in the initialization of the photometry pipeline and in the pipeline itself. We therefore introduce the TAOS implementation of these algorithms here before discussing the photometry initialization and pipeline.

FPAP has been used extensively in the past (see Da Costa 1992; Mighell 1999 and references therein). We have devised a method of FPAP using square apertures. Square apertures were chosen because of computational efficiency. A comparison test was performed to measure the difference between the TAOS method of using square apertures and the more standard circular aperture used in the IRAF package *apphot*. No difference in the signal-to-noise ratio (S/N) was detected. (Note that the use of circular apertures may give more accurate absolute photometry, however we are only concerned with the S/N values of our light curves. This is discussed further in § 6.)

The algorithm is illustrated in the left panel of Figure 3. In the diagram, pixels are labeled p_{ij}, where i is the row and j is the column. In this example, the aperture is centered at $i = 2.8$ and $j = 2.7$, and the aperture size is 1.8×1.8 pixels. When summing up the counts in each of the pixels, the counts are weighted by the fraction of the pixel area covered by the aperture. That is, the flux of a star is given by

$$F = \sum_{i}^{\text{rows}} \sum_{j}^{\text{cols}} f_{ij} c_{ij}, \tag{1}$$

where f_{ij} is the fraction of pixel p_{ij} covered by the aperture, c_{ij} is the photon count in pixel p_{ij}, and the sums are taken over the rows and columns covered by the aperture.

The TAOS FPAP method is also used to find the centroid positions of stars in different parts of the analysis pipeline. While a simple calculation of the barycenter of the point-spread function (PSF) will give an accurate estimate of the centroid in many cases, many of the stars in our images have significantly distorted PSFs, and the barycenter can be significantly offset from the actual peak of the flux distribution. We have found that the S/N measurements of the light curves are highest when the apertures are centered on the peak of the PSF rather than the barycenter. As will be discussed in § 4.2, the centroids of 50 bright reference stars are used to calculate the centroid position of the remaining stars in the image. The bright reference stars have aperture sizes of 5 to 7 pixels, and the barycenter can be offset from the peak by as much as 1 pixel. We thus use the TAOS FPAP method with a small aperture to find the peak of the flux distribution.

The method is an iterative algorithm, in which an offset to a starting centroid is calculated, the offset is added to the current centroid position, and the process is repeated until the offset approaches 0. To calculate the offset, we use a set of four 1.25×1.25 pixel apertures in a 2×2 grid, as shown in the right panel of Figure 3. The grid is centered on the expected centroid (x, y) of the star (the starting point depends on when in the process the algorithm is used, and will be described in the following sections), and the flux F_{ij} (where i and j correspond to the apertures shown in the right panel of Figure 3) in each aperture is calculated using the TAOS FPAP algorithm.

The offset to the centroid is then calculated by

$$\Delta x = \frac{-F_{11} + F_{12} - F_{21} + F_{22}}{F_{11} + F_{12} + F_{21} + F_{22}} \tag{2}$$

and

$$\Delta y = \frac{-F_{11} - F_{12} + F_{21} + F_{22}}{F_{11} + F_{12} + F_{21} + F_{22}}. \tag{3}$$

The offset is scaled by a factor ρ, chosen to improve the rate of convergence of the algorithm, and then added to the centroid, i.e.,

$$x \rightarrow x + \rho\Delta x,$$
$$y \rightarrow y + \rho\Delta y. \quad (4)$$

The process repeats until $\Delta x < 0.0002$ pixels and $\Delta y < 0.0002$ pixels.

A value of $\rho = 0.9$ is used because it was found to work well in most cases. If the algorithm fails to converge after 60 iterations, it is restarted with a value of $\rho = 0.45$. Once again, if the algorithm fails to converge after 60 iterations, it is restarted with a value of $\rho = 0.125$. If it fails to converge after 60 iterations, the star is ignored in further calculations using the centroids (see the following sections). Larger values of ρ typically lead to rapid convergence, while more iterations are needed with smaller values of ρ. However, in some cases with larger values of ρ, the offsets will start to oscillate between iterations without converging. A smaller value of ρ usually takes care of this problem. This algorithm works very well for the vast majority of the stars, and fails to converge after attempts using all three values of ρ less than 1% of the time.

3. INITIALIZATION OF THE PHOTOMETRY PIPELINE

Before running the photometry pipeline, we first generate a template aperture mask from a stare-mode image taken before the commencement of zipper-mode operation. The template is a list of aperture centroids and aperture sizes for all of the stars in the field down to a limiting magnitude of $R = 14$. This template is the basis of the aperture mask (see § 4.2) used in the exposure-based aperture photometry pipeline (see § 4). In addition, we choose 50 bright reference stars which are used to set the aperture centers of the remaining stars when the aperture mask is generated. These operations are discussed in the following subsections.

3.1. Template Aperture Mask Centroids

The template aperture mask centroids are determined from a 1 s stare-mode image taken before the start of zipper-mode operations. We run *SExtractor* (Bertin & Arnouts 1996) on the image and create a catalog of stars, which is a list of instrumental magnitude, centroid, and FWHM of each of the stars found in the image. Next, we run *imwcs*, from the *WCSTools* software package (Mink 2006). The *imwcs* utility uses the centroids listed in the *SExtractor* output in combination with the USNO-B1 catalog (Monet et al. 2003), to add World Coordinate System (WCS) keywords to the stare-mode image FITS file header. We then run *immatch*, another *WCSTools* utility, to match the stars in the *SExtractor* catalog with those in the USNO-B1 catalog. The *immatch* utility outputs a new catalog which includes magnitude, and the right ascension α and declination δ.

The pixel position of each star in *SExtractor* output is then used as the initial position to derive the center of the star by using the centroid-finding algorithm described in § 2. This recentering process is done to ensure we use the same centroid algorithm in the whole pipeline, in particular because the aperture size is determined based on these centroids (see § 3.3).

Next, we use a cubic polynomial transformation to fit the pixel coordinates (x, y) of stars to the celestial coordinates (α, δ) in the catalog. Stars for which the centroid-finding algorithm fails to converge are not used in the fit. We do the fitting in terms of the difference from the center of the field; that is, we define

$$\alpha_{\rm d} = \alpha - \alpha_{\rm c}$$
$$\delta_{\rm d} = \delta - \delta_{\rm c}$$
$$x_{\rm d} = x - x_{\rm c}$$
$$y_{\rm d} = y - y_{\rm c}, \quad (5)$$

where $(\alpha_{\rm d}, \delta_{\rm d})$ is the angular distance from the field center $(\alpha_{\rm c}, \delta_{\rm c})$, and $(x_{\rm d}, y_{\rm d})$ is the distance in pixels from the center of the CCD $(x_{\rm c}, y_{\rm c})$. We fit these coordinates to a set of polynomials of the form

$$x_{\rm d} = C_0 + C_1\delta_{\rm d} + C_2\alpha_{\rm d} + C_3\delta_{\rm d}^2 + C_4\delta_{\rm d}\alpha_{\rm d} + C_5\alpha_{\rm d}^2 + C_6\delta_{\rm d}^3 + C_7\delta_{\rm d}^2\alpha_{\rm d} + C_8\delta_{\rm d}\alpha_{\rm d}^2 + C_9\alpha_{\rm d}^3, \quad (6)$$

$$y_{\rm d} = D_0 + D_1\delta_{\rm d} + D_2\alpha_{\rm d} + D_3\delta_{\rm d}^2 + D_4\delta_{\rm d}\alpha_{\rm d} + D_5\alpha_{\rm d}^2 + D_6\delta_{\rm d}^3 + D_7\delta_{\rm d}^2\alpha_{\rm d} + D_8\delta_{\rm d}\alpha_{\rm d}^2 + D_9\alpha_{\rm d}^3. \quad (7)$$

The coefficients of the two-dimensional (2D) polynomial function are computed by Gauss-Jordan elimination (Press et al. 1994) with iterative 3σ clipping applied, with a maximum of three iterations. The root mean square of the residual between the observed and fit centroids of the reference stars is typically about 0.1 pixels.

We set the centroid positions in the template aperture mask to the centroids predicted by the polynomial fit. We have found that this provides a more accurate estimate of the centroid position than the values obtained from *SExtractor*, especially for the dimmer stars, where the centroid position will have a larger uncertainty.

3.2. Selection of Reference Stars

The next stage of the initialization is to select 50 reference stars from the template. These stars are used to calculate the differences between the centroids of the stars in each exposure and the template aperture mask centroids, in order to compensate for image motion. The selection criteria for these reference stars are that they are bright and isolated in the star field. For the brightnesses of the stars, we determine the instrumental

magnitude of each star in the stare-mode image using the FPAP algorithm described in § 2. We set the size of the aperture to twice the average FWHM of the stars listed in the *SExtractor* catalog (see § 3.1). A linear regression between the USNO-B1 R2 magnitude and the instrumental magnitude is used to derive the zero-point magnitude. The calibrated and instrumental magnitudes are saved for the further analysis of the light curves and are used in the calculation of the aperture size (see § 3.3).

The selection procedure is straightforward. We loop through from the brightest to dimmest stars in the list and then search for 50 isolated stars. The criterion for an isolated star is that the centroid distance to any star with $>10\%$ of its flux is greater than $5\sqrt{2}$ pixels. When calculating the centroid distance, we use y mod 76 rather than y in order to get the distance between the stars as they would appear in a row block rather than in the stare-mode image.

It is desirable to have these reference stars distributed across the field in order to enable an accurate transformation between the template and the aperture mask (see § 4.2). We have selected a set of 167 star fields for the TAOS survey (Lehner et al. 2009), and visual inspection of stare-mode images of all of these fields has confirmed that the brightest stars are uniformly distributed across the focal plane. Therefore we do not test for this in the photometry software.

3.3. Optimization of Aperture Size

The optimal aperture size for each star is determined empirically from the actual data. The TAOS aperture photometry pipeline is custom designed to run on a series of many nearly identical images, and we use this to our advantage by analyzing the S/N as a function of aperture size on the first few minutes of data in a zipper run. With this method we achieve the highest S/N across the entire light curve.

To find the aperture size which produces the highest S/N for each of the light curves, we loop over aperture sizes from a minimum of 1 pixel to a maximum $A_{\max}$ pixels for each star. We empirically choose the parameter $A_{\max}$ to be a function of the form

$$\begin{aligned} A_{\max} &= 1.9 + 57.5e^{-0.41(m+15)} \\ &\quad + (e^{-0.35(m+2)(a-1)} - 1) \quad \text{if} \quad A_{\max} < 12.5, \\ A_{\max} &= 12.5 \qquad\qquad\qquad\qquad \text{otherwise}, \end{aligned} \tag{8}$$

where m is the instrumental magnitude and a is the semimajor axis of the star. (The third term is used for elongated stars which arise from optical deficiencies near the corners of the field.) The semimajor axis a is computed by using the spatial moment function (Stobie 1980)[16]

$$a^2 = \frac{\bar{x}^2 + \bar{y}^2}{2} + \left[\left(\frac{\bar{x}^2 - \bar{y}^2}{2}\right)^2 + \overline{xy}^2\right]^{\frac{1}{2}}. \tag{9}$$

The moments $\bar{x}^2$, etc., are relative to the barycenter $(\bar{x}, \bar{y})$, and are calculated over a square region of pixels centered on the pixel corner nearest the centroid of the star, with width $\lfloor 1.75 \times \mathrm{FWHM} + 0.5 \rfloor$ pixels ($\lfloor\ \rfloor$ is the floor function, which returns the largest integer less than or equal to the argument). The FWHM value comes from the *SExtractor* catalog and the centroid from the centroid-finding algorithm (see § 3.1). Note that it is possible to get the value of a from the *Sextractor* output; however, we calculate this parameter when analyzing zipper-mode images as well (see § 5), so we use the same piece of code to calculate the parameter here.

We calculate the optimal aperture size for each star as follows. For each star, we loop over aperture sizes A from 1 to $A_{\max}$ in steps of 0.1 pixels. Using the aperture size A, we then generate the photometry aperture mask (see § 4.2) and run the TAOS photometry (see § 4) on the first 1000 zipper-mode exposures in the series. We then have a light curve of 1000 points for each star in the field. Next, we calculate the S/N for each light curve. This calculation is described in § 3.3.1.

Some example plots of S/N versus aperture size A are shown in Figure 4. The left panel of this figure shows S/N versus A for two isolated stars, one bright (top) and one dim (bottom). Note the oscillatory features in the S/N curves. Nearly every curve has a local maximum at aperture sizes of an integer $+0.7$, and a local minimum near integer aperture sizes. This is an artifact of our FPAP algorithm. In particular, the noise is found to increase with the rate that the total number of pixels in the aperture (either partially or completely) varies as the centroid shifts around. The rate at which pixels enter and exit the aperture is largest at integer aperture sizes. Another contributing factor is that the photoelectrons from a star are distributed among different numbers of pixels if the centroid is in the center of a pixel or near the boundary. This also tends to increase the noise levels when integer aperture sizes are used.

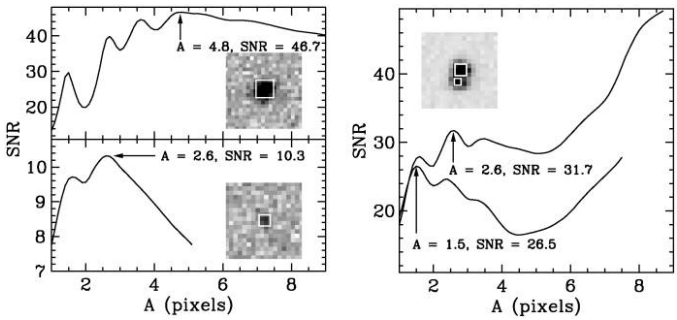

FIG. 4.—Plots of S/N vs. aperture size A for isolated bright (*top left*) and dim (*bottom left*) stars, and for two neighboring stars (*right*). Postage-stamp images (zipper-mode images after sky background and streak subtraction) with the optimal aperture sizes are shown for each star. For each star, the S/N curves run from $A = 1.0$ to $A_{\max}$.

[16] See also the *SExtractor* User's Manual.

In cases such as those shown in the left panels of Figure 4, the optimal aperture size is set to the value of A that gives the highest S/N. (In cases where there are two local maxima where the S/N values differ by less than 1%, we always choose the smaller aperture to minimize contamination from neighboring stars and streaks.) The situation is slightly more complicated in the case of two neighboring stars (Fig. 4, right). In both cases, each S/N curve reaches a local maximum, reaches a minimum at around $A = 5$, and then starts upward again, surpassing the local maximum found at the smaller aperture size. This upward turn at larger aperture sizes is due to contamination by the neighboring star. If the aperture gets large enough, it will cover both of the stars and give a higher S/N than that obtained for each individual star. We have found that if we require the maximum to be a local maximum (rather than the largest S/N value on the curve) and limit the possible aperture size to $A_{\max}$, we can prevent this from happening in nearly every case. The only time the additional criteria do not prevent the aperture from covering two stars is when they are so close as to be practically inseparable.

This algorithm works for most stars, but there are occasionally pathological cases where the optimum aperture size cannot be calculated. In some cases, the S/N curve increases monotonically, with no local maximum. This usually occurs for stars that show up in the stare-mode images but are too dim to be accurately measured in the zipper-mode image due to the increased sky background and reduced exposure time. In these cases we set the aperture size to $0.75 \times A_{\max}$ (these stars are generally not worth analyzing for the occultation search, but are nonetheless analyzed for possible future use). There are also cases with only one local maximum and one local minimum, and then a monotonic increase in the S/N curve. In such cases, the local maximum occurs at $A = 1.7$ and the minimum at $A = 2$,; the difference between the maximum and minimum are quite small. This usually occurs with stars near the corners of the focal plane that have badly distorted PSFs. In such cases we set the aperture size to $0.75 \times A_{\max}$. In other cases, the S/N growth curve cannot be classified by any of the above tests. This generally manifests itself in a monotonically decreasing S/N curve. In these cases, the best aperture size is set to 1 pixel. Any stars with such pathological S/N curves are flagged as such in the photometry output file.

3.3.1. Calculation of Signal-to-Noise Ratio

The simplest method to determine the S/N of a light curve is to calculate the mean μ and standard deviation σ over the entire light curve. However, TAOS light curves show low-frequency variations due to causes such as changing atmospheric transparency, and this can artificially raise the standard deviation over the length of the light curve relative to the local running mean. Furthermore, significant outliers (from cosmic rays, etc.) are not uncommon in the light curves, and these data can also corrupt the S/N calculation. We therefore have devised a method to calculate the S/N value after removing any slowly varying trends and significant outliers.

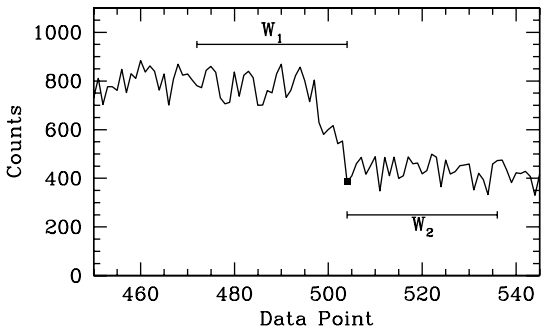

FIG. 5.—Simulation of a light curve with a significant trend. The point marked with the *filled square* is an outlier in window W_1, but not in window W_2.

The first step in the process is to calculate the trend. We do this by taking a rolling average over a window of 33 data points centered on each point in the light curve. That is, we define

$$\bar{I}_j = \frac{1}{33}\sum_i I_{j+i}, \tag{10}$$

where I_j is the intensity at point j in the light curve, and the sum is performed over a window of length 33 centered on the data point j. (Note that the window size is truncated at either end of the light curve.) For each point $\bar{I}_j$, we perform iterative 5σ clipping where the standard deviation σ is calculated over the same window. We chose to use 5σ clipping rather than the usual value of 3σ in order to remove true spurious points (cosmic rays, etc.), but not to remove the less significant outliers due to changing contamination from neighboring stars as the centroids move relative to each other (the noise we are concerned about when optimizing the aperture size).

To calculate the signal μ, we simply take an average of the smoothed light curve $\bar{I}$. Note that this is slightly different than simply averaging the original light curve with 5σ clipping, the difference being in the selection of outliers. In the case of a relatively fast varying trend, an outlier in one window may not be an outlier in another window, as illustrated in Figure 5.

To calculate the noise, we use the detrended light curve, that is, we define

$$I'_j = I_j - \bar{I}_j, \tag{11}$$

and calculate the standard deviation of I'. Once again, we perform iterative 5σ clipping on the standard deviation, as described.

4. EXPOSURE-BASED APERTURE PHOTOMETRY ON ZIPPER MODE IMAGE DATA

As mentioned in § 1.1, the TAOS aperture photometry is performed on the basis of exposure, rather than row block, in order

to handle time-dependent offsets and rotation of the focal plane relative to the star field. That is, all of the stellar flux collected at a single epoch is analyzed together. The stars analyzed are spread across 27 consecutive row blocks and make up a subset of the stars in each of these row blocks.

This is illustrated schematically in Figure 6. Figure 6a shows the original stare-mode image and the aperture mask template discussed in § 3. Figure 6b shows two zipper-mode FITS files placed end to end. Note that the star appearing on each row block boundary has photoelectrons that are read out in two separate row blocks, but with the row blocks placed end to end in series in the zipper-mode FITS file, all of the counts from that star appear in the FITS image as a single object. This is also true for the same star when it lies on the row block boundary corresponding to the separation between the two zipper-mode FITS files if the image data in each file are placed end to end.

Figure 6c shows the same two FITS files with the aperture mask applied to the first complete exposure in the first FITS file. (The stars are marked with solid black circles for clarity.) The mask is generated for each exposure from the template aperture mask; this process is described in § 4.2. At this point the FPAP algorithm is run on each aperture and the resulting stellar flux is written to disk. Figure 6d shows the aperture mask at the next exposure. Figure 6e shows the aperture mask spanning multiple FITS files (stars are indicated by solid black squares for clarity). Note that the top star spans the boundary between FITS files. Figure 6f shows the mask where a significant offset occurs (stars indicated by empty squares). This is the motivation for the exposure-based photometry, since the aperture mask is much easier to generate than would be possible in a row-block based analysis in cases like this.

When analyzing TAOS image data, the first zipper FITS file in the series is skipped. This is because there will be a gradient in the sky background until the first 27 row blocks are read out. The first row of the row block will have 105 ms of sky exposure. However, since the readout of a row takes about 1.25 ms, the second row will have 106.25 ms of sky background, and the final row will have nearly 200 ms of sky background. Rather than modify our sky background and streak subtraction algorithm (see § 4.1), it is simpler to simply throw out the first image, as this only comprises 32 row blocks out of a typical total of 27,000 in each data run.

The photometry pipeline consists of the following steps:

1. The second and third zipper FITS files are loaded and placed end to end in memory (i.e., in a single FITS image array).

2. Sky background and streaks are subtracted from the image data in memory (§ 4.1).

3. For each exposure, an aperture mask is generated (see § 4.2) from the template mask, the FPAP algorithm is applied to each aperture, and the photometric data are written to disk.

4. When every exposure comprising data in the first FITS image has been analyzed, the first image is dumped from memory and the next FITS file is loaded and placed end to end with

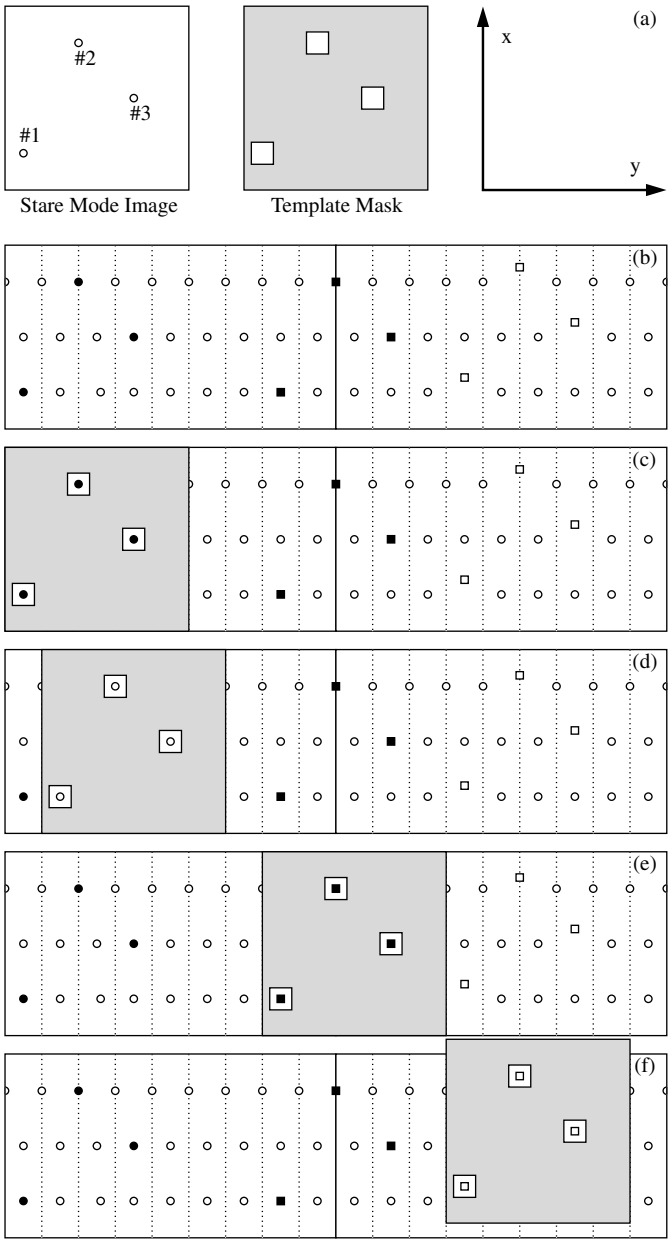

FIG. 6.—Illustration of TAOS aperture photometry. (*a*) Shows a stare-mode image of the field with three stars (*left*), the template aperture mask (*center*), and the pixel coordinate system (*right*). Time flows in the $+y$ direction in the lower panels. (*b*) Shows two zipper-mode FITS files stacked end to end. *Solid lines* indicates boundary between FITS files, and *dashed lines* indicate row block boundaries. Each FITS file contains 9 row blocks. (Note that actual TAOS FITS files contain 32 row blocks.) Stars are indicated by *open circles*. *Filled circles, filled squares,* and *open squares* indicate stars at three different epochs. Note that the star on the top spans row block boundaries. (*c*) Shows aperture mask applied to stars at first epoch (stars shown as *filled circles*), and (*d*) shows mask applied to next epoch The star spanning row block boundaries is treated as a single star when the mask is applied. (*e*) Shows mask applied across the boundary of the zipper-mode FITS files (stars shown as *filled squares*). The star spanning two files is treated as a single object, even though the flux from the star is in multiple FITS files. (*f*) Shows the mask applied with an offset (stars shown as *open squares*).

the image still in memory, as described in step 1. We wait until the first exposure comprising data only in the second FITS file has been analyzed, as image motion in the y direction may move some stellar flux into the previous FITS file.

5. Steps 2–4 are repeated for the remaining FITS files in the data run.

4.1. Streak and Background Subtraction

As discussed in § 1.1, stars will leave streaks in the images due to the finite readout time for each row. This can add noise to stars neighboring the streaks due to the time-dependent image motion in the zipper-mode images. This is illustrated in the right panel of Figure 2. Stars near the bright streaks will be contaminated with flux from streaks, and the amount of contamination varies with time as the centroids of the stars move relative to the streaks.

We have thus developed a simple algorithm to remove the streaks by subtracting background on a column-by-column basis for each row block. Note that this method subtracts sky background as well as any streaks in the images. For each column in a row block, the mode of the pixel values along the column is used as the sky background. The mode can be accurately estimated by (Da Costa 1992)

$$\text{mode} = 3 \times \text{median} - 2 \times \text{mean}, \tag{12}$$

where the median and mean are computed iteratively by applying 2.5σ clipping. Figure 7 shows the same image data from the right panel of Figure 2 after the application of the streak subtraction algorithm.

This simple method works quite well nearly all of the time. There are two situations, however, when it does not. The first case is when the tracking oscillations have frequencies $\gtrsim 5$ Hz. This is usually caused by gusty winds. In this case the streaks will not be straight lines, but will move back and forth across the columns (see Fig. 8, left panels). This will significantly increase the noise in any stars neighboring the bright streaks. The second case is when a slowly moving bright object moves across the field. An example of this effect is shown in the right panels of Figure 8. The object is moving slowly from left to right. The dark streak in the middle is the photons collected during the 105 ms exposure, and the two lighter streaks on either end of it are the streak collected during the readout. An algorithm to detect such situations automatically is under development.

4.2. Aperture Mask Generation

For each exposure run through the photometry pipeline, an aperture mask is generated from the template mask. The aperture sizes from the template are used in the aperture mask, but the centroid positions are recalculated to account for any image motion or rotation.

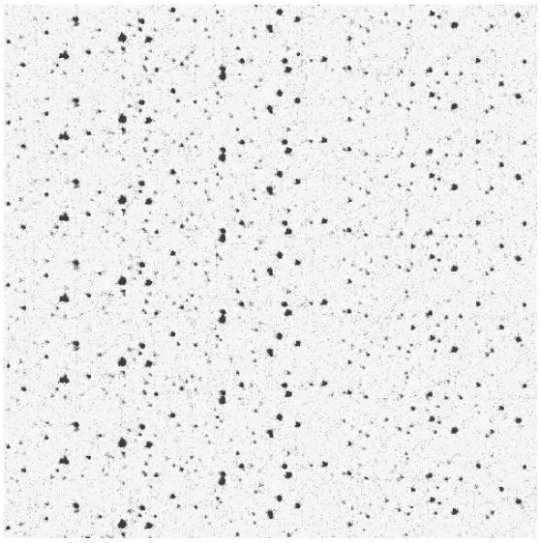

FIG. 7.—Same image data shown in the right panel of Fig. 2, after running the sky and streak removal algorithm.

To calculate the new centroids, the centroid-finding algorithm (see § 2) is run on each of the 50 bright reference stars selected during the photometry initialization phase (see § 3). When finding the centroids, it is important to have an accurate

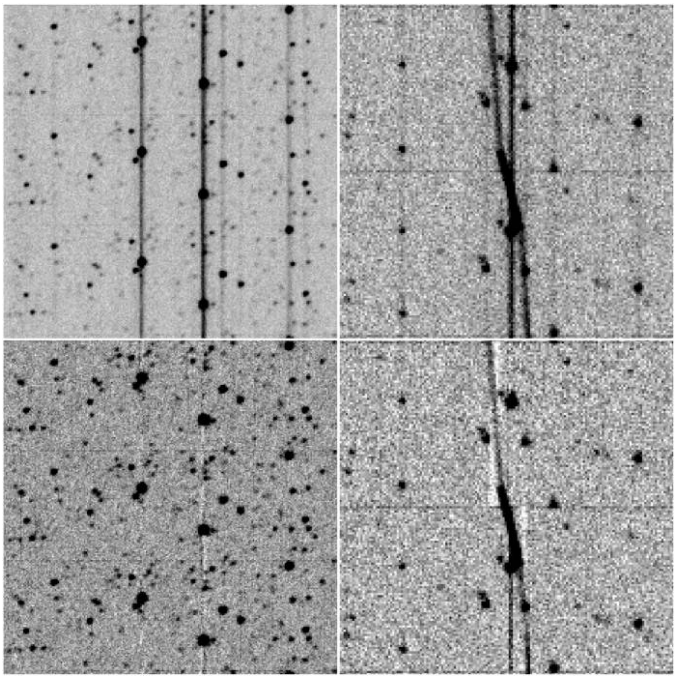

FIG. 8.—*Top left:* Series of row blocks collected under gusty conditions. *Bottom left:* Same as top left after sky and streak subtraction. Note the oversubtraction above and below the brightest star. *Top right:* Series of row blocks with a moving object. *Bottom right:* Same as top right, after sky and streak subtraction. Note the oversubtraction near the streak left by the moving object.

starting point for the algorithm, or it will never converge on the true centroids. To set a starting position, we begin with the centroid (x_k, y_k) for each of the 50 reference stars k from the previous exposure (for the first exposure we use the template centroids). While this is sufficient for the overwhelming majority of cases, situations occasionally arise where the centroid shifts by a pixel or more due to tracking errors, wind gusts, or other causes. Therefore, before running the centroid-finding algorithm, we run a test to determine whether the centroids have shifted by a pixel or more.

To run this test, we center a 10×10 pixel grid on the pixel corner nearest each centroid, and calculate the quantity

$$N_{ij} = \sum_{k=1}^{50} \frac{c_{k,ij}}{10^{-0.4m_k}} \tag{13}$$

for each pixel grid element $-5 \le i \le 4$ and $-5 \le j \le 4$, and with the sum taken over all 50 reference stars. Here m_k is the instrumental magnitude for the reference star k from the initial stare-mode image (see § 3), and $c_{k,ij}$ is the photon count in the pixel at

$$\begin{aligned} x &= \lfloor x_k + 0.5 \rfloor + i, \\ y &= \lfloor y_k + 0.5 \rfloor + j. \end{aligned} \tag{14}$$

N_{ij} is the average (weighted by the brightness of the reference star) over all reference stars k of the counts in the pixel, which is offset from each centroid (x_k, y_k) by a distance of (i, j). We then define $i_{\max}$ and $j_{\max}$ as the pixel grid element with the largest value of N_{ij}. The values of $i_{\max}$ and $j_{\max}$ indicate the average pixel offset from the centroids from the previous exposure. We then replace the starting positions of the centroid-finding algorithm with

$$\begin{aligned} x_k &\rightarrow x_k + i_{\max}, \\ y_k &\rightarrow y_k + j_{\max}. \end{aligned} \tag{15}$$

If there is no significant pointing offset, both $i_{\max}$ and $j_{\max}$ are equal to 0, and we use the centroids from the previous exposure.

At this stage, we run the centroid-finding algorithm and get a new centroid location (x_k, y_k) for each of the 50 reference stars. Next, we fit a 2D polynomial function to the template positions and current centroids of the reference stars. Note that any reference stars on which the centroid-finding algorithm fails to converge (see § 2) are not used in the fit. The polynomial function is of the form

$$x_k = C_0 + C_1 X_k + C_2 Y_k + C_3 X_k^2 + C_4 X_k Y_k + C_5 Y_k^2 \tag{16}$$

$$y_k = D_0 + D_1 X_k + D_2 Y_k + D_3 X_k^2 + D_4 X_k Y_k + D_5 Y_k^2, \tag{17}$$

where (X_k, Y_k) is the template centroid of star k. The coefficients of the 2D polynomial function are computed by Gauss-Jordan elimination (Press et al. 1994) with iterative 3σ clipping applied, with a maximum of three iterations. The root mean square of the deviations between the observed and fit centroids of the reference stars is typically about 0.1 pixels. (Note that we use a third-order polynomial in § 3.1 to convert from celestial to pixel coordinates, while we use a second-order polynomial here to translate the template centroids to the aperture mask centroids. This is because higher-order terms are needed for an accurate translation from celestial coordinates than from the template centroids.)

The fit is then applied to calculate the aperture mask positions for all of the stars in the template. At this point the aperture mask is complete, and the FPAP algorithm can be run on all of the stars in the current exposure.

5. DATA QUALITY PARAMETERS AND FLAGS

A set of data quality parameters and flags is calculated for each measurement of each light curve. These values are used as warnings during further analysis that the measurements are suspect, or in some cases, invalid. For each star we calculate the semimajor axis a, the semiminor axis b, and the position angle θ. The parameter a is calculated as shown in § 3.3, and the other parameters are calculated by

$$b^2 = \frac{\bar{x}^2 + \bar{y}^2}{2} - \left[\left(\frac{\bar{x}^2 - \bar{y}^2}{2} \right)^2 + \overline{xy}^2 \right]^{\frac{1}{2}}, \tag{18}$$

$$\theta = \frac{1}{2} \arctan\left(\frac{2\overline{xy}}{\bar{x}^2 - \bar{y}^2} \right), \tag{19}$$

where the moments $\bar{x}^2$, etc., are calculated as discussed in § 3.3.

The flags are defined as follows.

PHOT_OK: This flag simply means that the data point is not suspect.

PHOT_EDGESTAR: This flag is set if the centroid of the star is close enough to the image boundary that the aperture is not entirely within the focal plane.

PHOT_INTTIME: This flag indicates that the integration time for the current exposure is off by more the 5% from the nominal value. The readout cadence is controlled by the system clock on the camera control computer. On occasion the system will hang after an interrupt is received by some random process and this will cause two or three consecutive measurements to have the incorrect integration time.

PHOT_REFCOUNT: This flag is set if the number of reference stars for which a valid centroid is found during the aperture mask generation is less than 17.

PHOT_XOFFSET: This flag is set if the rms of the fit for the x component of the centroid is >1 pixel after the final (third) iteration of 3σ clipping (see § 4.2).

PHOT_YOFFSET: This is the same as the PHOT_XOFFSET flag, except for the y component of the centroid.

PHOT_BADFIT: This flag indicates that the polynomial fit to the centroids of the reference stars (see § 4.2) fails. This flag is also set if the number of reference stars is less than 12 (the number of parameters in the fit).

6. COMPARISON WITH SEXTRACTOR AND DAOPHOT

We have run *SExtractor* and *DAOPHOT* on a series of consecutive zipper-mode FITS files to test the performance of the TAOS photometry. For *SExtractor*, we investigated the various methods of running the software and found that the method using automatic apertures resulted in the highest S/N values. For *DAOPHOT*, we used the *allstar* task to do PSF fitting on the stars in the zipper-mode images. The results of the comparisons are shown in Figures 9 and 10. Figure 9 shows the light curves from all three photometry packages and the residuals to the TAOS light curves. Figure 10 shows scatter plots of the photometric intensities measured by *SExtractor* and *DAOPHOT* relative to those calculated from the TAOS photometry. A comparison of mean, noise, and S/N are shown in Table 1, as well as computing time. (S/N values are calculated on the detrended light curves, as discussed in § 3.3.1.)

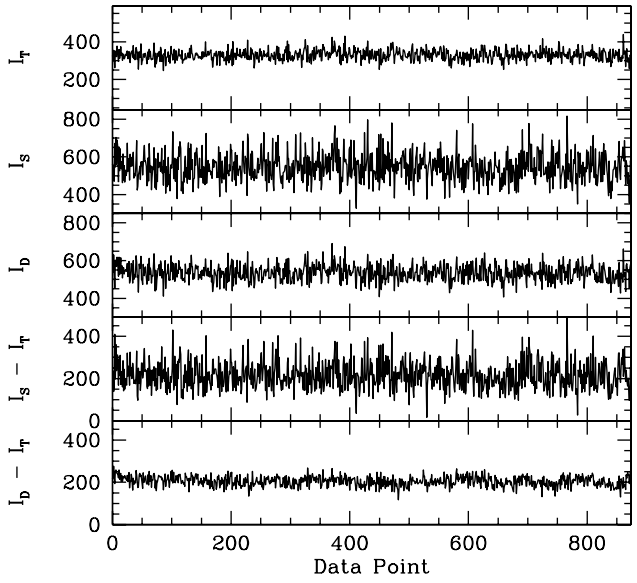

FIG. 9.—Comparison of *SExtractor* and *DAOPHOT* with the TAOS photometry. *Top panel:* Light curve for a $R = 11.7$ star from TAOS photometry (I_T). *Second panel:* Light curve from same star measured by *SExtractor* (I_S). *Third panel:* Light curve from same star measured by *DAOPHOT* (I_D). *Fourth panel:* Residual between TAOS and *SExtractor* light curves. *Bottom panel:* Residual between TAOS and *DAOPHOT* light curves. The lengths of the y-axes are the same for the top three plots in order to illustrate the relative noise characteristics of the three light curves. For all plots, only the first 165 s of the light curves are shown.

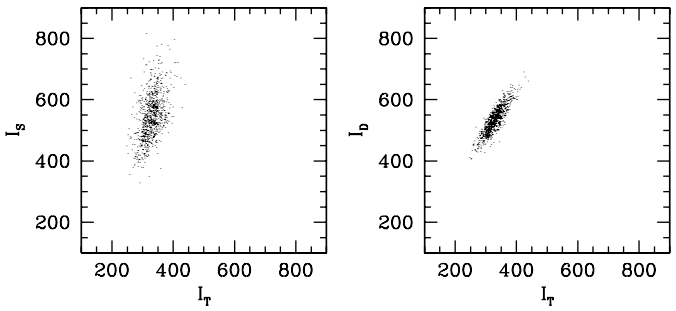

FIG. 10.—Scatter plots showing comparison of TAOS photometry to *SExtractor* (*left*) and *DAOPHOT* (*right*), for the same light curve data shown in Fig. 9. (Axes are all set to span the same range in order to accurately illustrate the relative responses of the various software packages.) The large variation on the photometric measurements is due to slowly varying trends in the data (see Zhang et al. 2008 and the discussion in § 3.3.1).

Several features are evident in the plots in Figures 9 and 10. First, the TAOS light curve in Figure 9 is an average of more than 200 counts lower than the light curves produced by *SExtractor* and *DAOPHOT*. This is because *SExtractor* and *DAOPHOT* are designed to perform accurate absolute photometry, whereas the TAOS software is optimized for S/N. The TAOS photometry finds the highest S/N values with an aperture smaller than would be needed to measure all of the flux in the star. Second, the light curve from *SExtractor* is relatively noisy. On the other hand, the *DAOPHOT* light curve is much less noisy than the *SExtractor* light curve, but still slightly more noisy than the TAOS light curve. In terms of S/N, it is shown in Table 1 that the TAOS photometry does nearly as well as *DAOPHOT*, and significantly better than *SExtractor*. Finally, there is little correlation between the TAOS and *SExtractor* measurements in Figure 10. However, there is a strong linear relation between the data measured by the TAOS photometry and *DAOPHOT*. A fit to a straight line through the scatter plot in the right panel of Figure 10 has a slope of about 1.6, which is nearly the ratio between the mean values of the two light curves, indicating that the sensitivity of the TAOS photometry to variations in the flux is nearly the same as the sensitivity achieved by *DAOPHOT*.

TABLE 1
COMPARISON OF TAOS PHOTOMETRY WITH SEXTRACTOR AND DAOPHOT

	TAOS	SExtractor	DAOPHOT
Mean Counts	329.9	545.9	535.2
Mean Residual from TAOS	.....	215.9	205.3
Noise	27.4	73.7	42.3
Residual Noise	.....	10.7	8.7
S/N	12.0	7.4	12.7
Computational efficiency[a]	1	7.9	0.0056

[a] Defined as inverse of the computational time needed to analyze a zipper-mode data run relative to that needed by the TAOS photometry.

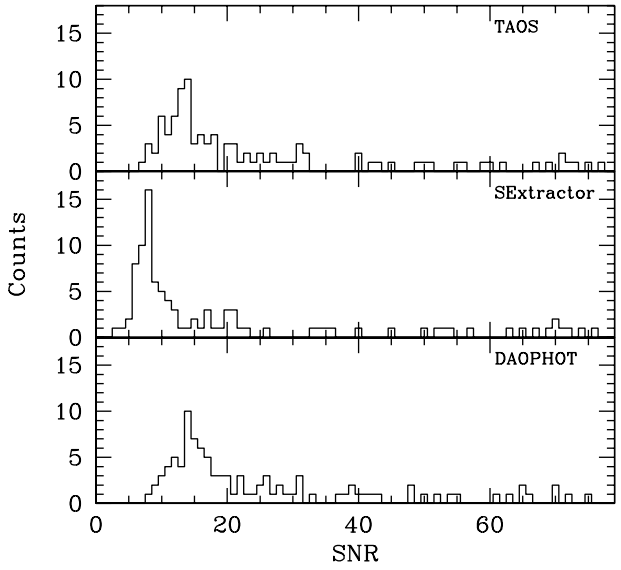

FIG. 11.—Histograms of S/N values from the TAOS photometry, *SExtractor*, and *DAOPHOT*. The same 100 light curves with the highest S/N from a single zipper-mode data run are included in each of the histograms.

While Figures 9 and 10 show light curve data from a single star, it is useful to see how the different packages perform on an entire set of stars. Figure 11 shows histograms of S/N values from the 100 stars with the highest S/N in a single zipper run. Note that TAOS performs comparably to *DAOPHOT*, while significantly outperforming *SExtractor*.

In terms of computational efficiency, *SExtractor* is the fastest, as it is capable of analysis at a rate of 224 row blocks s^{-1}, compared with a rate of 27 row blocks s^{-1} for the TAOS photometry. *DAOPHOT* is the slowest, capable of analysis at a rate of 0.16 row blocks s^{-1}. TAOS also has 240 s of overhead at the photometry initialization. This includes 200 s to collect the first 1000 exposures used to calculate the optimal aperture sizes and 40 s to generate the template aperture mask (using a single 2 GHz Pentium Xeon processor). The software then requires another 55 s to catch up with the data, meaning that real-time analysis commences after the first 1500 exposures, or 300 s of zipper-mode data.

7. CONCLUSION

We have developed a high-speed, computationally efficient, aperture photometry pipeline. The software produces light curves with S/N values nearly as good as those achieved with *DAOPHOT*, and significantly outperforms *SExtractor*. The software runs nearly 200 times as fast as *DAOPHOT*, and while slower than *SExtractor*, the TAOS photometry is fast enough to perform real-time photometric reduction on TAOS data, after an initialization period of about 300 s.

The work at National Central University was supported by grant NSC 96-2112-M-008-024-MY3. Y. I. B. acknowledges the support of National Research Foundation of Korea through grant 2009-0075376. Work at Academia Sinica was supported in part by the thematic research program AS-88-TP-A02. Work at the Harvard College Observatory was supported in part by the National Science Foundation under grant AST-0501681 and by NASA under grant NNG04G113G. S. L. M.'s work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory in part under contract W-7405-Eng-48 and by Stanford Linear Accelerator Center under contract DE-AC02-76SF00515. K. H. C.'s work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory in part under contract W-7405-Eng-48 and in part under contract DE-AC52-07NA27344.

REFERENCES

Bailey, M. E. 1976, Nature, 259, 290
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Da Costa, G. S. 1992, in ASP Conf. Ser. 23, Astronomical CCD Observing and Reduction Techniques, ed. S. B. Howell, 90
Lehner, M. J., et al. 2009, PASP, 121, 138
Mighell, K. J. 1999, in ASP Conf. Ser. 189, Precision CCD Photometry, ed. E. R. Craine, D. L. Crawford, & R. A. Tucker 50
Mink, D. 2006, in ASP Conf. Ser. 351, Astronomical Data Analysis Software and Systems XV, ed. C. Gabriel, C. Arviset, D. Ponz, & S. Enrique, 204
Monet, D. G., et al. 2003, AJ, 125, 984
Press, W. H., et al. 1994, Numerical Recipes in C (New York: Cambridge University Press)
Roques, F., Moncuquet, M., & Sicardy, B. 1987, AJ, 93, 1549
Roques, F., et al. 2009, Earth Moon and Planets, 105, 201
Stetson, P. B. 1987, PASP, 99, 191
Stobie, R. S. 1980, J. British Interplanet. Soc., 33, 323
Zhang, Z.-W., et al. 2008, ApJ, 685, L157